**תרגול 8** – 04.05.2010

## משפחות *Hashing* ב"ת בזוגות ומשפט *Valiant-Vazirani*

תזכורת:

קבוצה של מ"מ $X_1,\ldots,X_n$ שמקבלים ערכים מעל $\{1,\ldots,p\}$ נקראים ***k*-ב"ת** אם כל $k$ מהם ב"ת.

ניתן לבנות באופן מפורש התפלגות של $n$ משתנים $k$-ב"ת בגודל $O\left(n^{\left\lfloor\frac{k}{2}\right\rfloor}\right)$.

הבנייה עבור $k=2$: כדי לבנות $n-1$ מ"מ מעל $\{0,1\}$ שהם ב"ת בזוגות, ניתן להשתמש בכל הפולינומים הלינאריים הלא-טריביאליים מעל $\mathbb{Z}_2^{\log n}$. לוקחים מ"מ $Y_1,\ldots,Y_{\log n}$, ואז מגדירים לפיהם את $X_1,\ldots,X_n$. למשל, אם $X_1=Y_1+Y_3+Y_7$, $X_2=Y_1+Y_4+Y_{16}$, אזי $X_1=X_2$ אם ורק אם $Y_1+Y_4+Y_{16}=Y_1+Y_3+Y_7$, ומהבנייה והתכונות של $\mathbb{Z}_2$ נקבל את הרצוי.

### משפחות גיבוב ב"ת בזוגות

משפט: יהיו $m,n\in\mathbb{N}$, $m<n$. אזי קיימת משפחה $H_m^n$ של פונקציות מ-$\{0,1\}^n$ ל-$\{0,1\}^m$ המקיימת את הדרישות הבאות:

1. לכל $x,y\in\{0,1\}^n$ $x\neq y$ ולכל $a,b\in\{0,1\}^m$, $\Pr_{h\in_U H_m^n}[h(x)=a\wedge h(y)=b]=2^{-2m}$.
2. הקידוד של כל $h\in H_m^n$ הוא באורך $\mathrm{Poly}(n)$ ביטים.
3. בהינתן קידוד של $h\in H_m^n$ ונקודה $x\in\{0,1\}^n$, ניתן לחשב את $h(x)$ בזמן פולינומי באופן דטרמיניסטי.

משפחה כזאת נקראת **משפחת גיבוב ב"ת בזוגות**.

משפט *Valiant-Vazirani*: נניח שקיים אלגוריתם דטרמיניסטי פולינומי $A$, שעל קלט פסוק CNF $\varphi$, במידה ול-$\varphi$ יש *השמה מספקת יחידה*, הוא מוצא השמה זאת (אין שום הנחה לגבי פסוקים עם 0 או לפחות 2 השמות מספקות). אזי $\mathrm{NP}=\mathrm{RP}$.

הוכחה:

טענה הסתברותית: תהי $H_k^n$ משפחת גיבוב ב"ת בזוגות מ-$\{0,1\}^n$ ל-$\{0,1\}^k$ כמו קודם. תהי $S\subseteq\{0,1\}^n$ כך ש-$2^{k-2}\le|S|\le2^{k-1}$. אזי $\Pr_{h\in_U H_k^n}\left[\left|\{x\in S\,|\,h(x)=0^k\}\right|=1\right]\ge\frac{1}{8}$.

הוכחה:

יהי $N$ המ"מ שמייצג את מספר האיברים מ-$S$ שממופים ל-$0^k$. אזי $E[N]=2^{-k}\cdot|S|\in\left[\frac{1}{4},\frac{1}{2}\right]$. ההסתברות ש-$N\ge2$ היא לכל היותר $\binom{|S|}{2}\cdot\underbrace{2^{-2k}}_{\text{אי תלות בזוגות}}$.

ההסתברות ש-$N\ge1$, לפי עקרון ההכלה וההפרדה היא לפחות:

$$\sum_{x\in S}\Pr[h(x)=0^k]-\sum_{x<x'\in S}\Pr[h(x)=0^k\wedge h(x')=0^k]=$$

$$=2^{-k}\cdot|S|-\underbrace{\binom{|S|}{2}\cdot2^{-2k}}_{\text{אי תלות בזוגות}}\ge2^{-k}\cdot|S|-2^{-2k}\cdot|S|^2=2^{-k}\cdot|S|\cdot\left(1-2^{-k}\cdot|S|\right)\ge\frac{1}{8}$$

נראה אלגוריתם שבהסתברות לפחות $\frac{1}{8n}$, בהינתן פסוק CNF $\varphi$, נכריע האם $\varphi$ ספיק או לא (כמובן, בהינתן האלגוריתם $A$). האלגוריתם יקבל פסוק ספיק בהסתברות $\frac{1}{8n}$, ידחה פסוק לא ספיק בהסתברות 1.

**האלגוריתם**: (בהינתן פסוק $\varphi$ על $n$ משתנים)

ננחש באופן אקראי $2\le k\le n+1$ (מעכשיו, נניח שיש ל-$\varphi$ בין $2^{k-2}$ ל-$2^{k-1}$ השמות מספקות).

נבנה פסוק $\psi$ באופן הבא: $\psi(x)=\varphi(x)\wedge h(x)=0^k$. ניתן לתאר דרישה זו היות ו-$h$ ניתנת לחישוב יעיל, כאשר $h$ נבחרת באופן אקראי מתוך $H_k^n$.

**נכונות**:

נבחן 2 מקרים:

אם ל- $\varphi$ אין השמה מספקת, אז היות ו- $\psi=\varphi\wedge(\text{משהו})$ , אז גם ל- $\psi$ אין השמה מספקת.

אם ל- $\varphi$ יש בין $2^{k-1}$ ל- $2^{k-2}$ השמות מספקות, אז בהסתברות $\frac{1}{n}$ נבחר באלגוריתם את הטווח הנכון. נסמן ב- $S$ את קבוצת ההשמות הטובות ל- $\varphi$ . אזי מהטענה, בהסתברות לפחות $\frac{1}{8}$ יהיה בדיוק $x\in S$ יחיד עבורו $h$ מקרי שנבחר מקיים $h(x)=0$ . לכן, בהסתברות לפחות $\frac{1}{8n}$ בחרנו טווח נכון, וגם בחרנו $h$ שמקיימת שיש איבר יחיד שמקיים $h(x)=0$ . לכן, הרצת האלגוריתם $A$ תמצא את ההשמה, ולפיכך נכריע את SAT .